MERGING LANDSAT DERIVED LAND COVERS INTO QUAD-REFERENCED GEOGRAPHIC INFORMATION SYSTEMS

James M. White
Robert M. Ragan
University of Maryland, Remote Sensing Systems Laboratory
College Park, Maryland 20742

and

K. Peter Lade
Department of Sociology and Anthropology, Salisbury State College
Salisbury, Maryland 21801

There have been enough successful experiments and field applications to conclude that Landsat digital data is sufficiently accurate to define the land cover distributions required as inputs to regional planning and resource allocation models. Even though computer aided translation of raw Landsat data is extremely efficient, the adoption of this new technology by counties and other regional governements has been limited. A major problem continues to center on the difficulties of merging Landsat derived land covers into a geographic information system (GIS) that a regional government may have been using for a number of years. Typically, the data stored in such a GIS is referenced to USGS quadrangle sheets and/or state plane coordinates.

The paper describes an approach for merging multi-scene Landsat data bases into existing geographic information systems having 5-second or smaller cells. The approach uses the output from the State of Maryland's UNIVAC 1180-based Landsat classification program ASTEP (Algorithm Simulation Test and Evaluation) developed by NASA. The structure of the technique was designed to address the problems that emerged as part of the Landsat classification of the 64,000 square mile Chesapeake watershed involving twelve scenes that was conducted by the senior author as part of an EPA study. The paper describes the removal of overlap among adjacent scenes, the crossreferencing of ground control points, and the isolation of the appropriate pixels from the Landsat data base for subsequent positioning into a file containing ancillary data referenced to a specific USGS 7½ minute quadrangle sheet. Examples illustrate the clustering of classified Landsat pixels to define the dominant land use for each of 8,100 cells within a series of quadrangle sheets distributed over the State of Maryland.

The approach uses a hard copy terminal tied to an ASTEP algorithm through telephone lines. A coordinate digitizing board for inputing the position of ground control points is also valuable, although manual measurements are possible. The approach is quite efficient and should be especially attractive for use on regional scale studies.

1. INTRODUCTION

Many states, counties, public utilities and other organizations concerned with planning and managementon a regional scle have integrated computer-based geographic information systems (GIS) into their decision making processes. Properly designed and operated systems allow the decision maker to define the spatial distribution of current conditions within the area of interest and, in an increasing number of cases, conditions in surrounding areas. Of equal importance, a good GIS also allows the decision maker to better understand how the region evolved to its current state and to interpret trends that indicate future conditions. When relatively large areas are involved, the use of GIS and computer technologies are pivotal in the development of effective planning and management strategies.

Current and past land cover distributions are key elements in the GIS. Unfortunately, these land cover files are often poorly defined or not up-to-date because of the times and costs required to assemble the data, interpret it, encode it and then enter into the GIS. Professionals concerned with GIS have long recognized the potential of digital format data from the Landsat series of satellites as a base for maintaining up-to-date land cover files in their systems. Although there are many successful applications, Landsat has remained a "potential" to the typical GIS user because the data format is less than ideal. This is especially true for grid cell based GIS that are referenced to USGS or state plane coordinate systems. Individual cells in such systems typically run from north-south vectors and may be 10, 91.8 or 4.5 acres or they may be 5 seconds in size. Although there are a number of programs designed to geometrically correct and reformat Landsat data, the time required to learn these systems, the level of effort and often the special equipment required limits the widespread application of many of the techniques and, thereby, Landsat continues to be unrealized potential.

2. THE CHESAPEAKE BAY EXPERIMENT

The development of the land cover distributions of the 64,000 square mile Chesapeake Bay watershed can be used to illustrate some of the problems that regional planning and management organizations encounter when attempting to integrate a Landsat derived data base into their operations. Figure 1 shows the outline of the Chesapeake watershed and the geometry of the 12 scenes used. The objectives of the Chesapeake Bay Project were: 1) produce a Level I land cover classification of the Chesapeake Bay watershed; 2) within agriculture land cover, determine tillage practices; and 3) tabulate land cover statistics by river subbasins. The land cover statistics were required as input to a mathematical model to predict the non-point source pollution loads to the Chesapeake Bay. The classification was conducted by the Northern Virginia Planning District Commission for the Environmental Protection Agency and used the IDIMS (Interactive Digital Image Manipulation System) and GES (Geographic Entry System) at NASA's Goddard Space Flight Center. The scenes had the known geometric distortions corrected (deskewing, removal of synthetic pixels) and procedures were developed to remove the overlap among scenes. The result was a properly registered land cover distribution that, through the use of a digitizer, was summarized for 63 subwatersheds distributed throughout the basin.

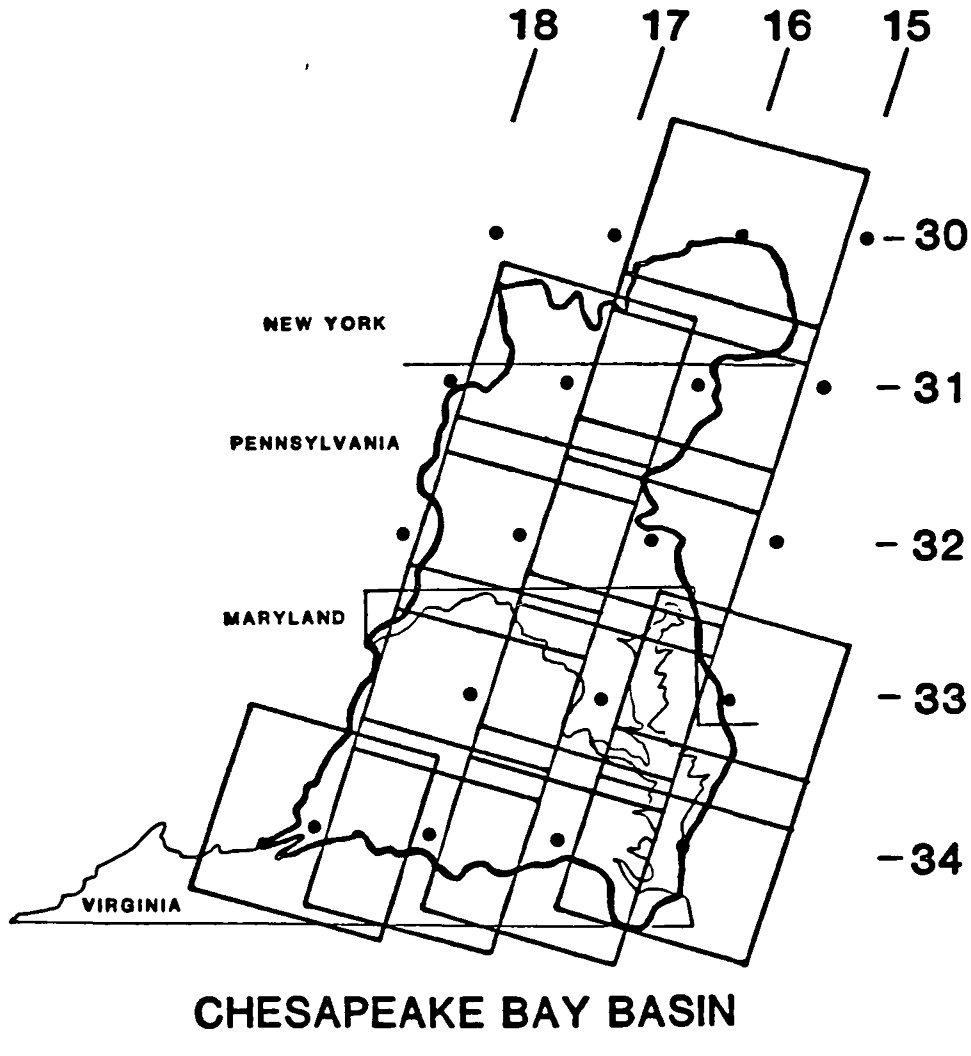


FIGURE 1

Landsat Scene Location within Chesapeake Bay Basin

The existence of such a digital data base was, obviously, very attractive to organizations within the area that had computer-based GIS. State systems, such as Maryland's MAGI (Maryland Automated Geographic Information System) and county systems, such as MSDAMP (Multi-Scale Data Analysis Mapping Program) used by Montgomery County, Maryland appear to be the logical recipients of the data derived in the Chesapeake study. While it was straight forward to extract a relatively large polygon such as a watershed from the Chesapeake Landsat data base, MAGI and MSDAMP require the definition of land covers within individual cells referenced to USGS or state plane coordinates. The general concept of MSDAMP is illustrated in Figure 2. MSDAMP is a series of 90 x 90 five second cells referenced to 19 USGS 7½ minute quadrangle sheets. The user obtains information by entering the name of the quadrangle sheet or sheets of interest and then extracts information by defining either polygons or individual cells. Maryland's MAGI uses either a 91.8 or 4.54 acre cell. Geographical Information Systems of the MSDAMP and MAGI types must have one dominant land cover defined for each cell in the data base. A schematic of the definition problem is illustrated in Figure 3. The domain of a particular USGS quadrangle sheet must be isolated from the Landsat data base and then a specific five second cell must become computer retrievable to the staff of the user organization. Image processing capabilities are available to all Maryland state and local governmental organizations through the State's UNIVAC 1100 series of computers located at the University and State College campuses. As potential State and county users of the Maryland portion of the Chesapeake data base moved toward integrating this additional information into their GIS, it became obvious that the efforts were not going to be widely successful because the needed software did not exist in a form that was compatible with UNIVAC 1100 series computers. There was no parallel software that could: rotate the Landsat coordinate system; reference the individual cells to USGS coordinates; isolate an array of cells defining a USGS 7½ minute quadrangle sheet and then resample the individual pixels to define a single land cover category for a predefined cell size. Without such software, the Chesapeake data base provided an excellent source of qualitative information, but remained inaccessable to the day-to-day user of the established computer-based geographical information systems operating within the State.

3. OBJECTIVES

If the Chesapeake Landsat-derived data base and similar future Landsat efforts are to be integrated into the existing geographical information systems, it is necessary to develop additional software to overcome the problems discussed above. To be useable, the additional software has to be fully integrated into established computer based approaches that are accessable and familiar to the users. Because few of the users in the State of Maryland have access to color CRT-based interactive image processing systems, the software had to be designed to run on a standard UNIVAC 1108 mainframe computer and require no more than a modem-connected hard copy terminal for operation. Further, because of severe restrictions placed on core storage during the daytime hours, the system had to be designed for minimum core storage utilization. With these contraints in mind, system development was undertaken to meet the following objectives:

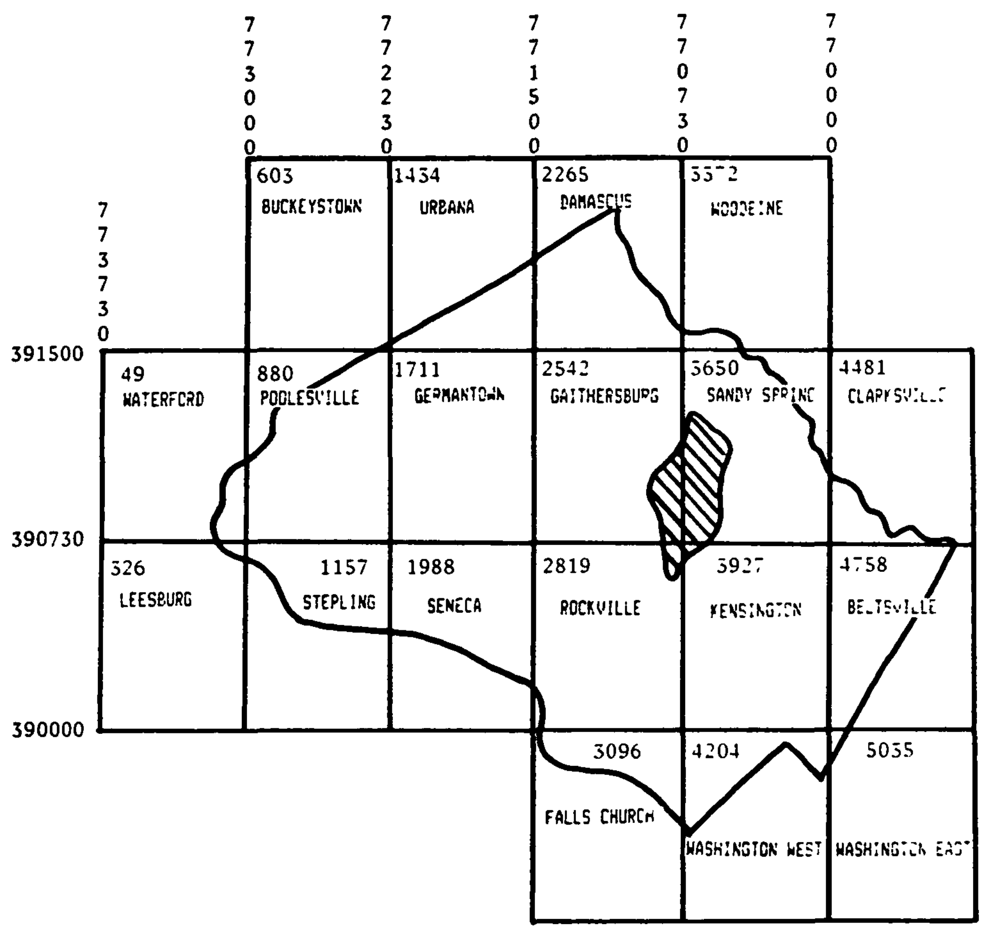


Figure 2

Quad-Sheet Storage Arrangement For
Montgomery County, Maryland

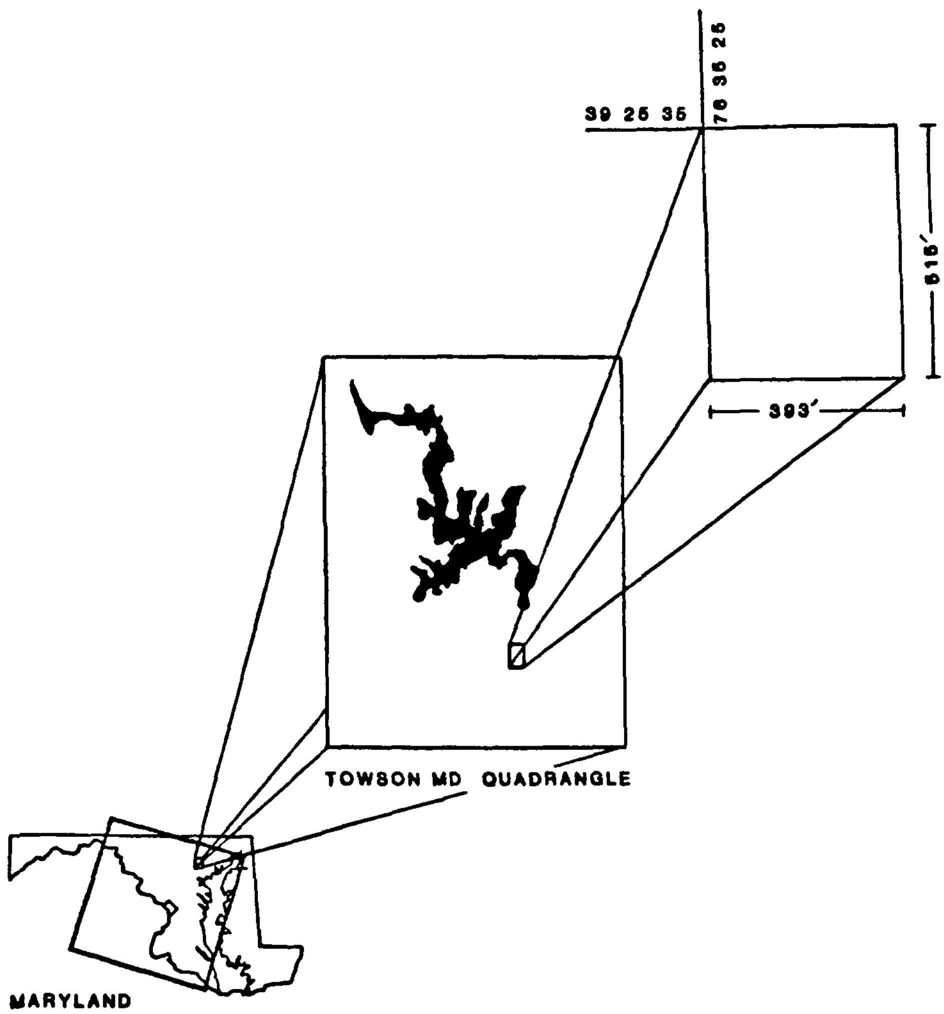


Figure 3

Isolation of a Single Cell from a Landsat Scene

1) Develop interactively an equation which relates a Landsat coordinate system to a latitude longitude coordinate system.

2) Create a transformed data base from Landsat imagery compatible with preconfigured Geographical Information Systems.

3) Enter geographic data from a map surface into ASTEP.

4. SYSTEM CAPABILITIES

To meet the objectives listed above, two programs (REGISTER and TRANSFORM) were developed. The program REGISTER was designed to input geographic data from a map surface and develop a regression relating the Landsat coordinate system to a latitude, longitude coordinate system. The program TRANSFORM, using the equations developed by the program REGISTER, was designed to create a geometrically correct data base compatible with preconfigured geographical information systems.

The program REGISTER serves two functions. First, it outputs to a file the longitude and latitude of points digitized from a map surface. Second, it provides an equation relating the latitude, longitude of a point to its Landsat line and sample coordinate. To complete the first function, a link (equation) must be developed relating the position on a map surface to its latitude and longitude. The position on the map surface may be input as coordinates from a digitizing table or measured manually off the map using the upper left corner as the origin. (Note: While manually measuring the location of points on a topographic map may be tedious, if it is done carefully, accuracy on a 1:24000 scale map can be ±40 feet.) The user inputs to the program are the longitude and latitude in degrees.minutes.seconds of the upper left corner of the map, the size of the map in minutes, the distance between "tic-marks" on the map in minutes, and the coordinates of the "tic-marks" from the digitizing table or as measured manually by the user. A first order polynomial regression equation is then developed relating the coordinate from the map surface to its longitude and latitude coordinate. A list of the actual and predicted coordinates, as well as the residuals, is produced for each of the "tic-marks" are output. The user has the option of removing any of the "tic-marks" from the registration if they were incorreclty digitized. The user also has the option of changing the regression equation to second or third order. (Note: For large scale maps, i.e., 1:24000 there should be no need to go to a second or third order equation.) Once the map has been registered to the digitizing table, the location of the ground control points can be digitized from the map. These points are then stored in a file for use in developing the transformation equation.

The second function that the program REGISTER performs is to allow the user to develop an equation relating the latitude, longitude coordinate system of the Geographical Information System to the line and sample coordinate of Landsat. The program reads the file containing the ground control points created above and a least-square file is applied to the points to develop a simple linear transformation of the form:

$$\hat{X} = C(1) + C(2)Y + C(3)X$$

$$\hat{Y} = C(9) + C(10)Y + C(11)X$$

where $\hat{X}$ and $\hat{Y}$ are the estimated sample and line values in the Landsat coordinate system, X and Y are the observed values of longitude and latitude (digitized coordinates) in the GIS coordinate system, and C(1), C(2)...C(N) are the coefficients fo the transformation equation expressed in the form:

$$[\hat{Y}\hat{X}] = [1YX] \quad \begin{matrix} C(1) & C(9) \\ C(2) & C(10) \\ C(3) & C(11) \end{matrix}$$

An output table is printed that contains the estimated sample and line value, the observed sample and line value, and the error (observed-estimated) sample and line value for each ground control point.

Upon examination of the ground control points, the user has the option of altering the list of ground control points. The user is prompted: DO YOU WISH TO EDIT POINTS? Y/N. If the user responds with an upper case Y he is prompted with: ADD(A) DELETE(D) OR EXIT(E)?. If the user wishes to delete a point, he responds with an upper case D. (Note: the development of the equation is an iterative process, the user may wish to restore a ground control point that was previously deleted by responding A.) The user is then prompted: INPUT NUMBER(S) TO BE ADDED OR DELETED ZERO (Ø) TO END. The user then would input the number(s) of the ground control point to be deleted, Ø indicates there are no more points. The user is then prompted: ADD(A) DELETE(D) OR EXIT (E)? and would respond E. The first order regression equation is then recalculated and the output table is again listed. The user has the option of editing points and recalculating the first order regression equation until he is satisfied that all the ground control point residuals have the same order of magnitude. When the prompts to edit points are answered N, the user will be prompted with: DO YOU WISH THIS TO BE THE HIGHEST ORDER? Y/N. If the response is N, a second order equation is developed with the form:

$$\hat{X} = C(1) + C(2)Y + C(3)X + C(4)Y^2 + C(5)X^2 + C(6)XY$$

$$\hat{Y} = C(9) + C(10)Y + C(11)X + C(12)Y^2 + C(13)X^2 + C(14)XY$$

The output table is printed and the user is given the option of editing points or developing a third order equation. The third order equation has the form:

$$\hat{X} = C(1) + C(2)Y + C(3)X + C(4)Y^2 + C(5)X^2 + C(6)XY + C(7)Y^3 + C(8)X^3$$

$$\hat{Y} = C(9) + C(10)Y + C(11)X + C(12)Y^2 + C(13)X^2 + C(14)XY + C(15)Y^3 + C(16)X^3$$

When the final transformation equation has been calculated, the equation is stored in a disc file for use by the program TRANSFORM.

The program TRANSFORM reads the transformation equation developed above and prompts the user for information concerning the location and size of the transformed area. The program prompts the user with: INPUT LONGITUDE, LATITUDE OF UPPER LEFT CORNER OF STUDY AREA IN DD.MMSS. When the user responds, he is prompted: INPUT THE SIZE OF TRANSFORMED AREA IN MINUTES LONGITUDE, LATITUDE. The user is not constrained to having the size of the transformed area the same in both longitude and latitude. The user is then prompted for the number of cells in the X and Y directions in the transformed area. The program TRANSFORM displays the Landsat sample and line value for the four corners of the transformed area and the minimum subset of the original data needed to transform the area. (Note: This allows the user to redefine the original study area to use the minimum amount of computer storage and CPU time.) The user then has the option of stopping the run to subset the original data, or continuing the run and creating the transformed area.

Figure 4 is a schematic representation of the procedure to transform a study and form a Landsat line and sample coordinate system to a latitude, longitude coordinate system. The output from the program TRANSFORM is a file that can be read directly into a geographic information system or reformatted by a program (INASTEP) to be entered back into ASTEP to use its statistical and map generating capabilities.

5. PROCEDURE

The procedure to transform a study area from a Landsat referenced coordinate system into a georeferenced coordinate system is as follows:

1) Ouput lineprint maps of study area

2) Locate and digitize features that can be found on both lineprint maps and topographic maps.

3) Develop regression equation

4) Transform the data.

The first step in transforming the data is to output lineprint maps from ASTEP, such as that illustrated in Figure 5, of the study are for use in locating features (ground control points). The lineprint map generation is the most critical portion of locating ground control points. A lineprint map is limited to displaying one channel of data with a practical limit of 20 grey levels, therefore, whatever a user can do to combine information from more than one MSS channel of data on a lineprint map is important. There are many ASTEP output products which are useful in the production of lineprint maps. A grey level map (density slice) of channel 7 can provide good land/water interface detection; it can also be useful in locating bridges, river boundaries, and power line clear cuts. A grey level map of channel 5 is useful in finding man-made features such as road intersections and industrial parks.

There are three ASTEP routines that allow the user to output information from more than one MSS channel of data. A map from the norm of all four channels (brightness map) can often be used to agument the output products

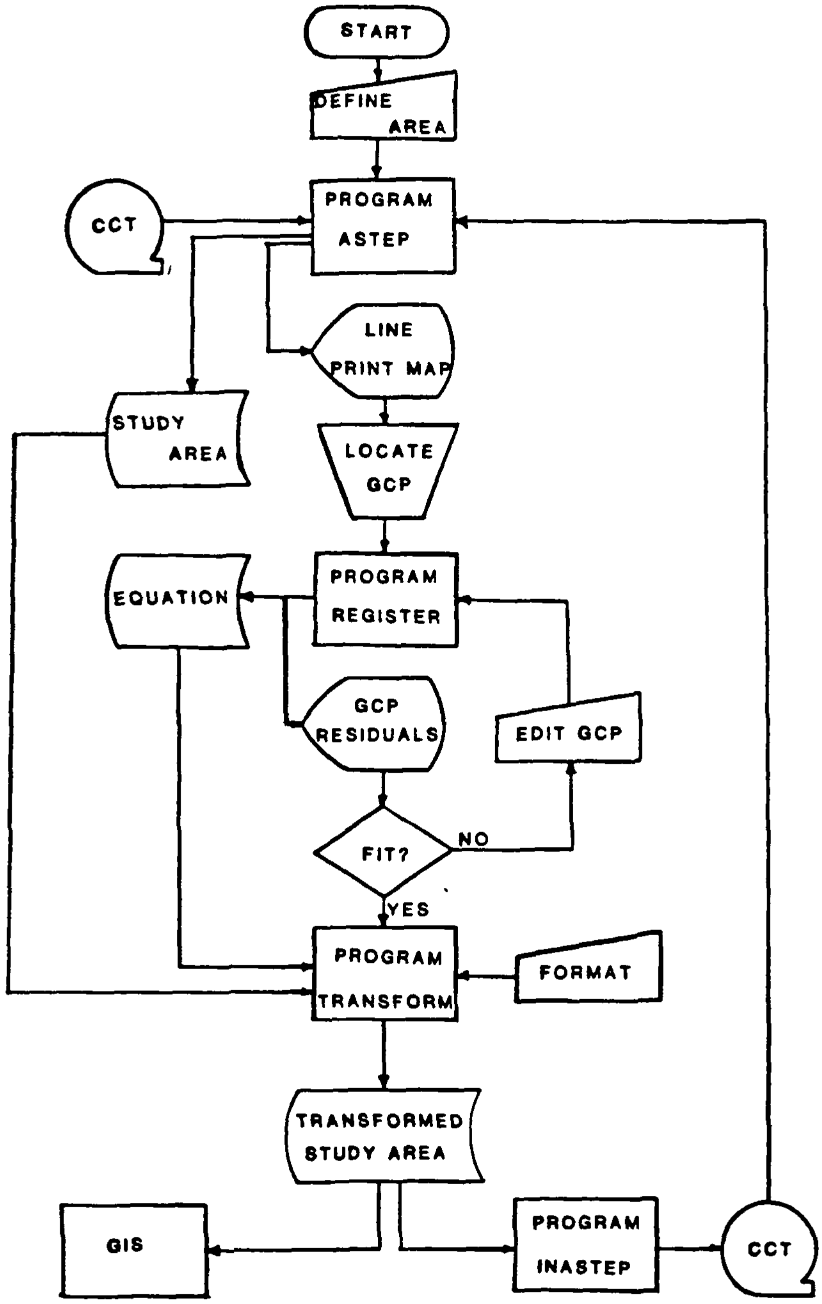


Figure 4

Procedure to Transform Landsat Data

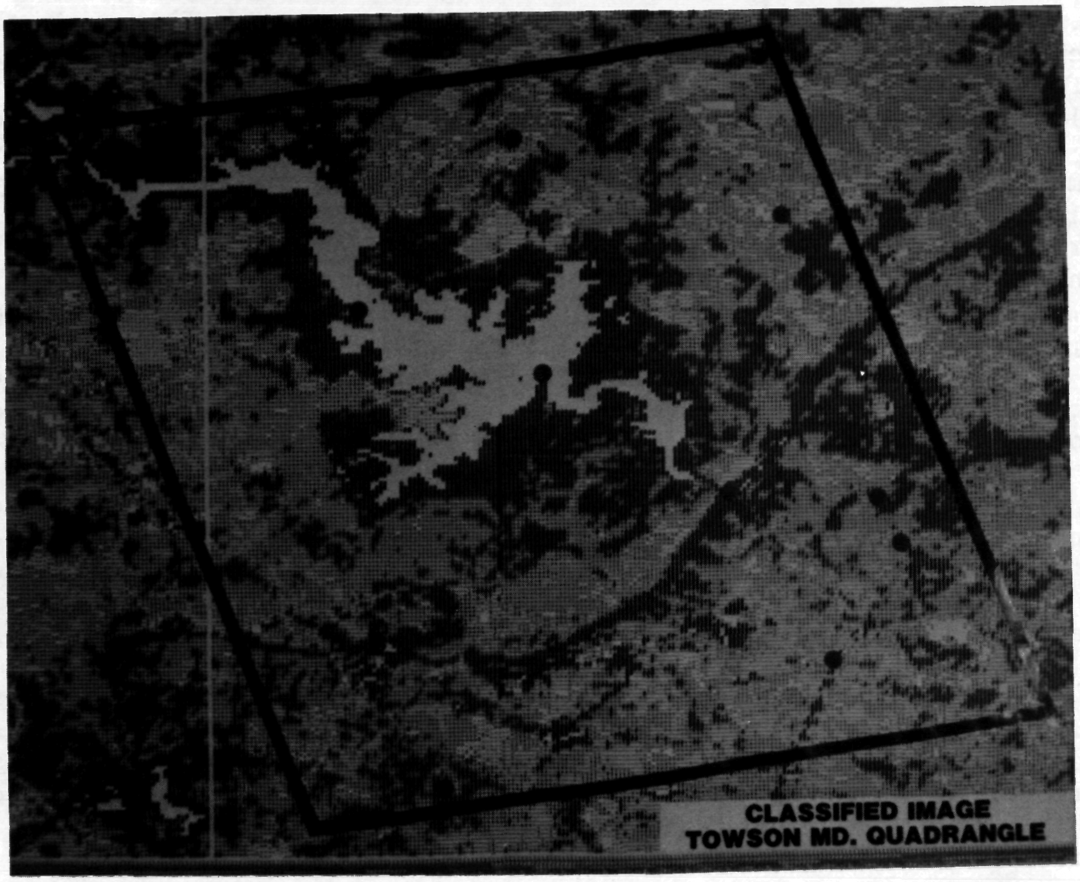


FIGURE 5

Unsupervised Classification of Study Area

listed above. An unsupervised classification of the study area using relatively few classes (6-8) provides a quick method of locating forest and grass field boundaries. Figure 5, for example, is an example of an unsupervised classification using only 6 classes.

Once the lineprint maps of the study area have been generated, the ground control point location can begin. It is important to find ground control points that are uniformly distributed throughout and surrounding the study area. A general rule of thumb is that in order to have confidence in the coefficients, there should be at least four ground control points for each of the coefficients of the regression equation. Therefore, a first order equation should have a minimum of 12 ground control points, a second order should have a minimum of 24 ground control points, and a third order should have a minimum of 32 ground control points. Depending on the size of the study area, the land cover, and the topography, it may not be feasible to find as many as 30 ground control points using lineprint maps. The ground control point can be any fixed feature locatable on both the lineprint maps and topographic maps. They may include bridges, islands, road intersections, power line clear cuts, and small ponds. It is generally easier to locate ground control point form images in early spring or late fall when there are no leaves on the trees to obscure ground features. There are some features however, that are easier to locate in summer scenes (i.e., power lines, roads).

When all the ground control points have been located, two files are created for each topographic map in the study area. The first file ocntains the digitized coordinates of the "tic-marks" on the topographic map. The second file contains the digitized coordinates and Landsat sample and line coordinates for each of the ground control points.

After the ground control points have been digitized, the process of developing the regression equation can begin. Tables I - V are examples from a program runstream which illustrates the process of developing a regression equation. For the sake of simplicity, only control points from one topographic map will be used. The user reponses are underlined and comments are in brackets.

TABLE I

Initial Output Used to Verify Regression Equation Defining Coordinates of Points on Quadrangle Sheet

@XQT RSSL*REGISTER.ABSTS

INPUT 1 TO DIGITIZE,2 TO DEVLOP REGRESSION,0 TO QUIT

1

ENTER LONGITUDE AND LATITUDE OF THE UPPER LEFT CORNER
SIZE OF TOPO IN MINUTS(LON,LAT),DISTANCE BETWEEN THE
REGISTRATION POINTS IN MINUTS(LON,LAT)

76.3730 39.3000 7.5 7.5 2.5 2.5

INPUT RERISTRATION POINTS

@ADD TRY1. [TRY1. is the file that contains the digitized coordinates of the "tic-marks"]

#	X/	Y/	X	Y	EX	EY
1	76.37300	39.30001	76.37300	39.30000	-.0000004	-.0000084
2	76.35001	39.30000	76.35000	39.30000	-.0000082	-.0000002
3	76.32300	39.30000	76.32300	39.30000	-.0000059	.0000004
4	76.30001	39.30000	76.30000	39.30000	-.0000086	-.0000012
5	76.37299	39.27299	76.37300	39.27300	.0000121	.0000064
6	76.34599	39.27299	76.35000	39.27300	.0000066	.0000070
7	76.32300	39.27300	76.32300	39.27300	.0000016	.0000035
8	76.30000	39.27300	76.30000	39.27300	-.0000012	.0000000
9	76.37300	39.24600	76.37300	39.25000	.0000039	.0000018
10	76.34599	39.25000	76.35000	39.25000	.0000066	-.0000018
11	76.32298	39.25000	76.32300	39.25000	.0000164	.0000006
12	76.29599	39.25000	76.30000	39.25000	.0000113	-.0000010
13	76.37302	39.22300	76.37300	39.22300	-.0000219	-.0000051
14	76.35000	39.22299	76.35000	39.22300	-.0000016	.0000053
15	76.32301	39.22301	76.32300	39.22300	-.0000094	-.0000141
16	76.29600	39.22299	76.30000	39.22300	.0000008	.0000082

#	X COEFF	Y COEFF
1	276326.42283975470368000000	141730.89427985332736000000
2	.01966168255291389490	19.76698345728208928000
3	-25.46664236535343654400	-.00409724894996088550

ERROR SQ = .18980407714843749984
SUM ERR X= .023438 SUM ERR Y = .015625
DO YOU WISH TO EDIT POINTS? Y/N

N

DO YOU WISH THIS TO BE THE HIGHEST ORDER? Y/N

Y

INPUT GROUND CONTROL POINTS # THEN DIGITIZED COORDINATES AND LABEL

@ADD TRY. [TRY. contains the digitized coordinates of the ground control points and their line and sample coordinates.]
@EOF

INPUT 1 TO DIGITIZE,2 TO DEVELOP REGRESSION,0 TO QUIT

Table I is a list of the predicted and actual longitude and latitude of the "tic-marks" on the topographic map, as well as the errors (actual-predicted).

where:

X1 = predicted longitude of "tic-mark"

Y1 = predicted latitude of "tic-mark"

X = actual longitude of "tic-mark"

Y = actual latitude of "tick-mark"

EX = X - X1

EY = Y - Y1

In this example, all the errors are less than 0.2 of a second (approximately 16' at this latitude) so there was no need to edit points or increase the order of the regression equation. After the user replies Y to the prompt "DO YOU WISH THIS TO BE THE HIGHEST ORDER?" the user is prompted for the ground control points and digitized coordinates.

The program uses the equations generated in Table I to convert the digitized coordinates of the ground control points to their corresponding longitude, latitude and stores the results for later use. The process is repeated for each topographic map in the study area. After all maps have been digitized, the user responds "2" to the prompt "INPUT 1 TO DIGITIZE, 2 TO DEVELOP REGRESSION, 0 TO QUIT".

TABLE II

Output Used to Verify Regression Equation

2

#	X/	Y/	X	Y	EX	EY
1	1598.32462	125.30842	1600.00000	126.00000	1.6753845	.6915751
2	1700.71982	121.36526	1707.00000	121.00000	6.2801819	-.3652592
3	1752.95477	102.56207	1767.00000	103.00000	14.0452271	.4379263
4	1673.43864	162.95376	1670.00000	164.00000	-3.4386444	1.0462418
5	1766.56570	142.39142	1772.00000	143.00000	5.4342957	.6085815
6	1800.19225	214.15857	1794.00000	212.00000	-6.1922455	-2.1585655
7	1749.11740	170.16807	1749.00000	170.00000	-.1174011	-.1680679
8	1811.80373	223.93128	1804.00000	223.00000	-7.8037262	-.9312840
9	1793.64261	169.85357	1795.00000	169.00000	1.3573914	-.8535690
10	1693.76706	247.05843	1674.00000	245.00000	-19.7670593	-2.0584335
11	1730.37199	283.31932	1775.00000	290.00000	44.6280060	6.6806755
12	1799.00040	254.53150	1783.00000	252.00000	-16.0003967	-2.5314999
13	1764.99077	209.58345	1758.00000	209.00000	-6.9907684	-.5834541
14	1718.71126	179.44728	1714.00000	180.00000	-4.7112579	.5527229
15	1678.39890	195.36758	1670.00000	195.00000	-8.3988953	-.3675785

#	X COEFF	Y COEFF
1	133544.1496423408385600000	38163.4309860281644800000
2	-.2120317696066082995 2	-.3870392918463467756 8
3	-.369058888613324142 08	.061604733148342915 92

ERROR SQ = 3226.5967561784200288 0000

SUM ERR X= .000092 SUM ERR Y = .000011

The program reads the file containing the latitude, longitude and line sample values for each ground control point and develops a regression equation relating latitude, longitude to line and sampple. Table II is a list of the actual line and sample, predicted line and sample and the errors for each of the ground control points. Where:

X1 = predict sample value

Y1 = predict line value

X = actual sample value

Y = actual line value

EX = X = X1

EY = Y - X1

The user is prompted "DO YOU WISH TO EDIT POINTS? Y/N". Deciding which point(s) to remove from the regression equation is somewhat of an art. A good rule-of-thumb would be to remove any point whose errors are significantly different from the rest (i.e., point 11). Being an iterative process, the user can delete points to see the effects and later add them if he wishes. In this example, the point "11" is deleted.

TABLE III

Output Used to Verify Regression Equation Without Point #11

```
DO YOU WISH TO EDIT POINTS? Y/N

Y

ADD(A) DELETE(D) OR EXIT(E) ?

D

INPUT NUMBER(S) TO BE ADDED OR DELETED ZERO(0) TO END

11

0

ADD(A) DELETE(D) OR EXIT(E) ?

E
```

#	X/	Y/	X	Y	EX	EY
1	1597.25696	125.22393	1600.00000	126.00000	2.7430420	.7760677
2	1706.68803	122.19216	1707.00000	121.00000	.3119659	-1.1921606
3	1765.75558	104.30714	1767.00000	103.00000	1.2444153	-1.3071394
4	1669.70222	162.44365	1670.00000	164.00000	.2977753	1.5563526
5	1772.50732	143.16842	1772.00000	143.00000	-.5073242	-.1684227
6	1794.33221	213.26280	1794.00000	212.00000	-.3322144	-1.2628002
7	1748.62469	170.05817	1749.00000	170.00000	.3753052	-.0581665
8	1804.76753	222.86302	1804.00000	223.00000	-.7675323	.1369820
9	1795.93893	170.10218	1795.00000	169.00000	-.9389343	-1.1021824
10	1675.03110	244.43800	1674.00000	245.00000	-1.0310974	.5619984
12	1785.26913	252.53663	1783.00000	252.00000	-2.2691345	-.5366306
13	1757.85748	208.53445	1758.00000	209.00000	.1425171	.4655514
14	1714.56311	178.84788	1714.00000	180.00000	-.5631104	1.1521187
15	1668.70555	194.02156	1670.00000	195.00000	1.2944489	.9784431

#	X COEFF	Y COEFF
1	134280.8078772947187200000	38173.8998919324948480000
2	-.1502657478089162163	-.3782592884064812627
3	-.4035720730124978656	.0570406292279734825

```
ERROR SQ =          31.582919040462002048
SUM ERR X=     .000122  SUM ERR Y =     .000011
```

Table III lists the output for the regression equation developed without ground control point "11". There are no ground control points having errors significantly different from the rest, so the user responds N to the prompt "DO YOU WISH TO EDIT POINTS? Y/N". The user is then prompted with "DO YOU WISH THIS TO BE THE HIGHEST ORDER? Y/N". If the user is not satisfied with the size of the errors, he will respond "N" and a second order equation will be developed, as illustrated in Table IV.

TABLE IV

Output to Verify Regression Equation Using Second Order Equation

```
DO YOU WISH TO EDIT POINTS? Y/N

N

DO YOU WISH THIS TO BE THE HIGHEST ORDER? Y/N

N

 #       X/           Y/          X            Y              EX          EY
 1  1599.11740   126.62065  1600.00000   126.00000      .8825989   -.6206522
 2  1707.99907   122.20280  1707.00000   121.00000     -.9990692  -1.2027960
 3  1767.02563   103.49714  1767.00000   103.00000     -.0256348   -.4971447
 4  1670.40628   163.15906  1670.00000   164.00000     -.4062805    .8409386
 5  1772.93092   142.48143  1772.00000   143.00000     -.9309235    .5185738
 6  1793.28535   212.64375  1794.00000   212.00000      .7146454   -.6437511
 7  1748.69167   169.81061  1749.00000   170.00000      .3083344    .1893864
 8  1803.47786   222.15224  1804.00000   223.00000      .5221405    .8477650
 9  1795.66638   169.25774  1795.00000   169.00000     -.6663818   -.2577400
10  1674.15500   245.56256  1674.00000   245.00000     -.1549988   -.5625591
12  1783.61646   252.21816  1783.00000   252.00000     -.6164551   -.2181587
13  1757.14552   208.36650  1758.00000   209.00000      .8544769    .6335011
14  1714.68178   179.08262  1714.00000   180.00000     -.6817780    .9173832
15  1668.79240   194.93917  1670.00000   195.00000     1.2075958    .0608253
  #                     X COEFF                          Y COEFF
  1        55196.0729053318499840000        8554.4074491299688448000
  2            -.7365065438243618816             .1563540830716192424O
  3             .4667681910814280838 4          -.0088954560706042684 3
  4             .0000024584913521641 4          -.0000008096422351028 3
  5            -.0000014794254793344 7           .0000004145494031438 7
  6            -.0000003765799286153 9          -.0000011060387239726 6
   ERROR SQ =           13.0905001915525644 8000
 SUM ERR X=     .008270  SUM ERR Y =    .005571
 DO YOU WISH TO EDIT POINTS? Y/N

N

 DO YOU WISH THIS TO BE THE HIGHEST ORDER? Y/N

Y

  END PROGRAM REGISTER.
```

The output shown in Table IV is again listed for the second order equation and the user has the option to edit points or go on to a third order equation. Once the user responds "N" to each question, the program creates a file with the coefficients of the regression equation.

After the regression equation has been developed, the user can transform the study area. The program TRANSFORM using the regression equation developed above, prompts the user for location and format of the transformed area. The program reads the raw Landsat data from a file created by ASTEP, transforms the data and outputs the transformed data in a format compatible with various geographic information systems. The user will be prompted for the longitude, latitude of the upper left corner of the study area, and the size of study area in minutes. The user is not required to have the study area correspond to one topographic map, and the study area can have different dimensions in the latitude and longitude direction. Table V is an example run of the program TRANSFORM; the study area is the Towson, MD quadrangle. (See Figure 6) The output file is to have the data stored in 5 second cells.

TABLE V

Example Row for Towson, MD

```
 INPUT LON,LAT OF UPPER LEFT CORNER OF AREA IN D.MS
76.3730 39.3000
 INPUT SIZE OF STUDY AREA IN MINUTS LON,LAT
7.5 7.5
 INPUT NUMBER OF CELLS LON,LAT IN TRANSFORMED AREA
90 90
 LANDSAT COORDINATES OF TOPO SHEET

 1590., 121.**********1770.,  93.
      I                     I
      I                     I
      I                     I
      I                     I
      I                     I
 1655., 292.**********1834., 264.


 IF YOU WISH TO SUBSET STARTING LINE STARTING SAMPLE# LINES # SAMPLES
                                88         1584        209     256


 DO YOU WISH TO SUBSET ? Y/N
Y
 DO YOU WISH TO QUIT ? Y/N
Y
 END PROGRAM TRANSFORM
```

FIGURE 6

USGS 7,5' Topographic Map Towson, MD

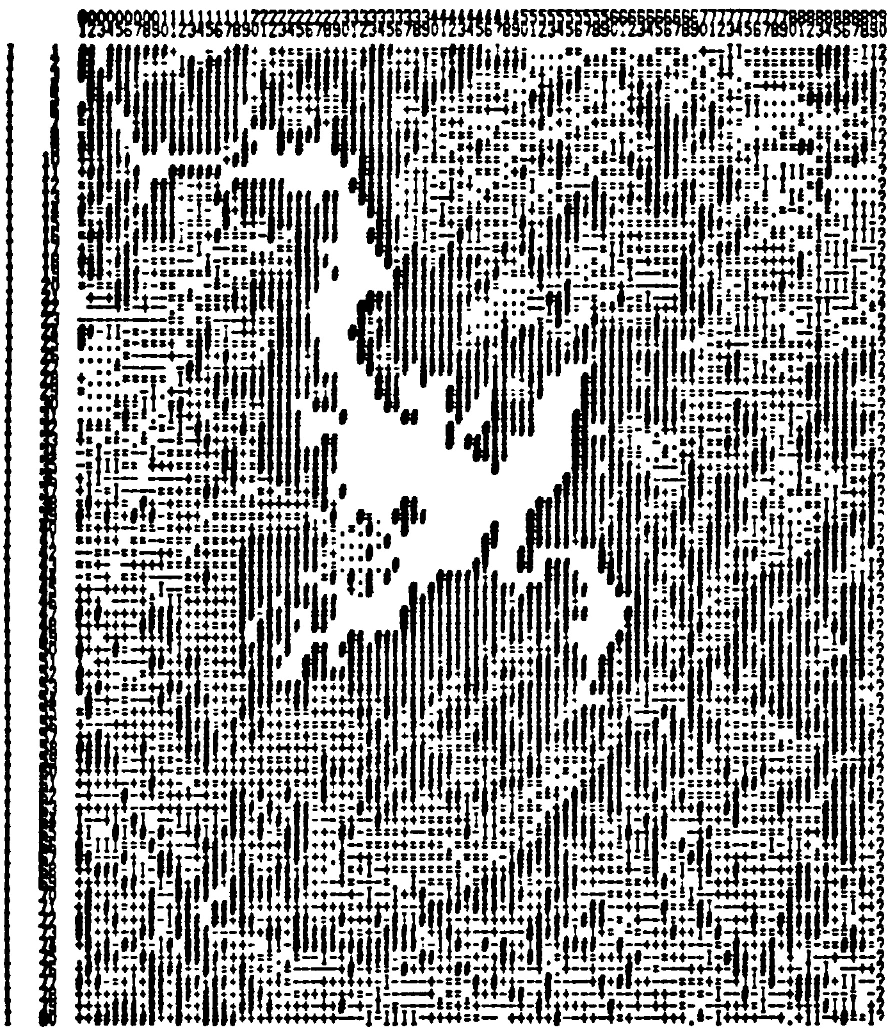
Figure 5 is the raw Landsat data used in the example in Table V, and Figure 7 is a map of the output file.

The user, by changing the number of cells in the output transformed area can produce output products at a given scale (i.e., 1:24000). If in the example above, had changed the number of cells in the transformed area from 90 x 90 to 212 x 272, a map of the output file would have a scale of 1:24000. Figure 8 is an example of such a product.

6. CONCLUSION

Many current or potential users of digital format remotely sensed imagery are restricted to the use of a remote lineprinter type terminal that accesses processing software on a general purpose, mainframe computer. The software described in the present paper was designed to provide this group of users with some of the interactive geometric corrections and data manipulation capabilities found on dedicated, color CRT-based image processing systems such as IDIMS. The system developed is compatible with ASTEP input/output reoutines and the UNIVAC 1100 series core limitations. It requires only a typewriter type terminal and is, therefore, available to Maryland State and local government users.

The interactive editing capabilities allow the user to produce a ±1 pixel registration accuracy between an image and map referenced position. Flexible output formate routines allow interfacing with preconfigured geographical information systems. With minor modifications, the system can easily be adapted to other geographical formats (i.e., state plane, UTM) and other sensors (i.e., RBV). The resulting transformed data bases can be re-entered into the ASTEP program to allow the user access to ASTEP capabilities such as scaled map production and statistical tabulations.



CLASSIFIED IMAGE

TOWSON MD. QUADRANGLE

GEO-CORRECTED 90x90 5 SECOND CELLS

Figure 7

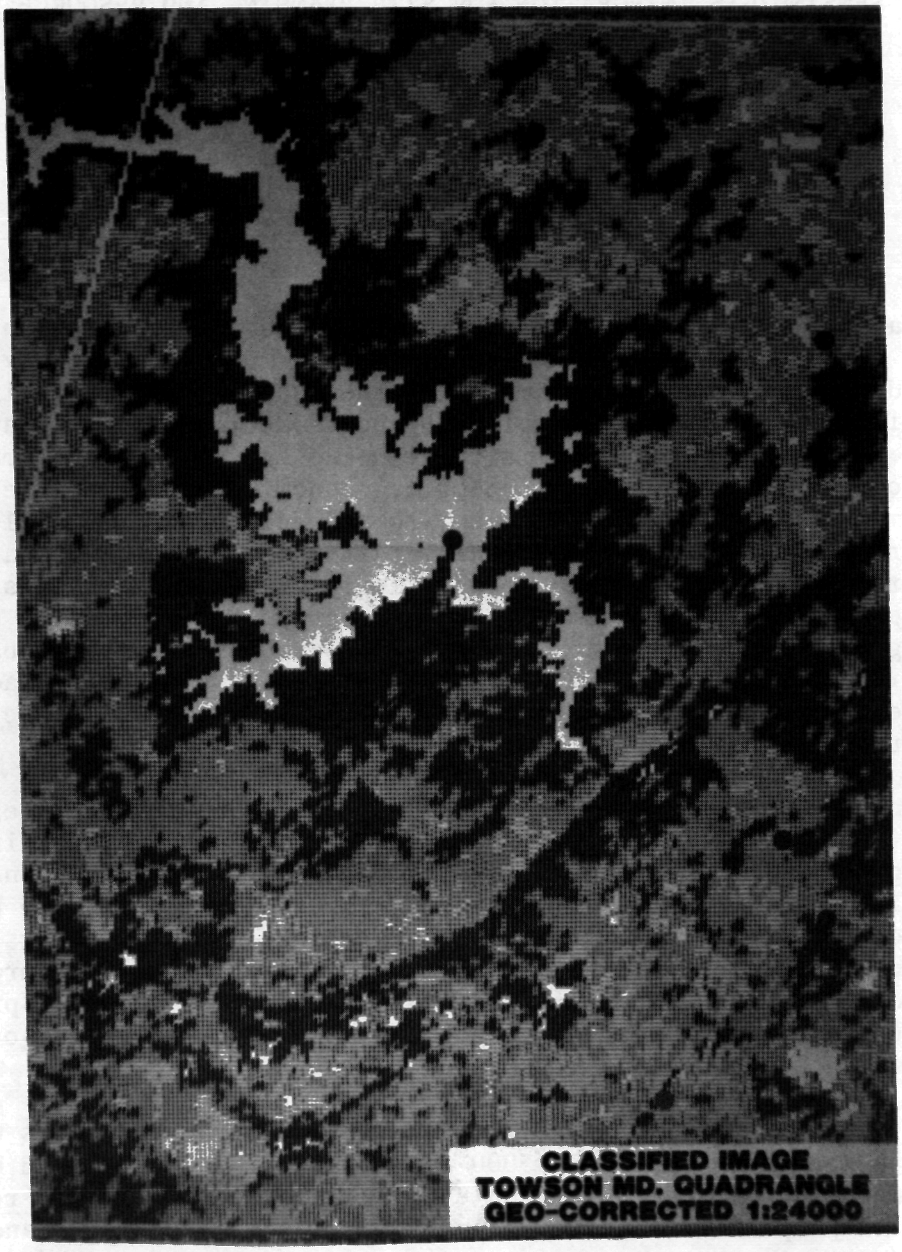


FIGURE 8

Quad-Centered Transformation Scale 1:24000